

SOCY 5021: DATA ANALYSIS
SPRING 2006
M 3-5:50; KTCH 33

Instructor: Professor Jason Boardman
Office: 206A Ketchum Hall
Office Hours: Monday 1-3

Email: boardman@colorado.edu
Phone: 303-735-0172
Fax: 303-492-8878

Texts

- 1) Allison, Paul D. 1999. *Multiple Regression: A Primer*. Thousand Oaks, CA: Pine Forge Press.
- 2) Hoffmann, John P. 2004. *Generalized Linear Models: An Applied Approach*. Boston: Pearson.
- 3) Kohler, Ulrich, and Frauke Kreuter. 2005. *Data Analysis Using Stata*. College Station, TX: Stata Press.

Objectives

This course will introduce students to multivariate statistical analysis and it will stress the application of these methods. I will assume that all students have had an elementary statistics course at the undergraduate level. Thus, I will review basic descriptive statistics and inferential statistics but we will transition into regression techniques by the third week of the course. The goal of the course is to teach students the skills and concepts that are necessary to perform and interpret elementary, intermediate, and some advanced statistical analyses in the Social Sciences. Accordingly, the primary project of the class is to write an original research paper using any of the techniques that we discuss in the class. We will work on these papers throughout the course.

The course is divided into three sections: 1) introduction to STATA; 2) regression analysis for continuous dependent variables; and 3) regression analysis for categorical dependent variables. The bulk of our time will be spent on the second section of the course in which we will thoroughly examine all aspects of the OLS regression model (estimation, assumptions, interpretation, and application). The last section of the class will extend these models to binary, multinomial, and count dependent variables.

Assessment

Your final grade in the class will be based on the following:

- 1) Participation (10%): this includes class attendance and participation.
- 2) Assignments (40%): assignments are typically made for each week of class involving the analyses of data for your project.
- 3) Final paper (50%): your paper must include-the following sections a) literature review; b) research questions/hypotheses; c) methods (data, measures, statistical model); d) findings; e) conclusions, f) references, and g) tables/figures.

Date	Topic
1/16	NO CLASS-Martin Luther King Day
1/23	<p style="text-align: center;">Data Analysis Using STATA</p> <p>Topics:</p> <ul style="list-style-type: none"> • What does a data set look like? How to read data into STATA? Managing files in STATA (.do, .log, .dta). • Descriptive statistics; Inferential statistics; Bivariate statistics • Sampling weights <p>Reading (to be read by class 1/23):</p> <ul style="list-style-type: none"> • Kohler & Kreuter Ch. 1, Ch 3 (pp. 45-59), Ch. 4, Ch. 5 (pp. 75-80). <p>Assignment (Due next Monday, 1/30):</p> <ul style="list-style-type: none"> • See handout and provide a <u>brief description of your research project (no more than 1-2 paragraphs).</u>
1/30	<p style="text-align: center;">Correlation and Bivariate Least Squares</p> <p>Topics:</p> <ul style="list-style-type: none"> • Model estimation/assumptions, parameter estimates, standard error, Type I error, one-tail vs. two-tail test <p>Reading:</p> <ul style="list-style-type: none"> • Kohler & Kreuter 177-192 • Allison 97-108 <p>Assignment (Due 2/13 in class):</p> <ul style="list-style-type: none"> • See handout.
2/6	<p style="text-align: center;">Visually exploring data: figures, graphs, and predicted values</p> <p>Topics:</p> <ul style="list-style-type: none"> • Predicted values, Scatter plots, Histograms, Q-Q plots, Box-whisker plots <p>Reading (to be read by class 1/30):</p> <ul style="list-style-type: none"> • Kohler & Kreuter Ch 6-7 <p>Assignment (Due 2/6 in class)</p> <ul style="list-style-type: none"> • See handout.
2/13	<p style="text-align: center;">Measuring and interpreting the independent variable</p> <p>Topics:</p> <ul style="list-style-type: none"> • Scale (NOIR), dummy variables, non-linearity and model fit, creating scales, Chronbach's alpha <p>Reading:</p> <ul style="list-style-type: none"> • Allison 156-159 • Bryman and Cramer "Constructing variables" * (CH2 Hardy and Bryman) <p>Assignment (Due 2/20 in class):</p> <ul style="list-style-type: none"> • See handout.
2/20	<p style="text-align: center;">Multiple Regression</p> <p>Topics</p> <ul style="list-style-type: none"> • The logic and estimation of the multiple regression model <p>Reading:</p> <ul style="list-style-type: none"> • Allison Ch 1, 2, 4 • Kohler & Kreuter 192-195 • Hoffman Ch 1,2. <p>Assignment (due 2/27 in class):</p> <ul style="list-style-type: none"> • See handout

2/27	<p style="text-align: center;">Assumptions and Problems</p> <p>Topics:</p> <ul style="list-style-type: none"> • Regression diagnostics <p>Reading:</p> <ul style="list-style-type: none"> • Allison Ch. 3, 6, 7 • Hoffmann pp. 10-17 • Kohler & Kreuter 199-216 <p>Assignment (due 3/6 in class):</p> <ul style="list-style-type: none"> • <u>Draft of the literature review for the paper.</u> Roughly 5 pages (double spaced). This should have a minimum of 5 papers (although it is not uncommon to cite 10-20 papers in a literature review).
3/6	<p style="text-align: center;">Model Building and Interaction terms</p> <p>Topics:</p> <ul style="list-style-type: none"> • Stepwise regression, standardized coefficients, path analysis, using the NESTREG command in STATA, seemingly unrelated estimation techniques (SUEST), substantive examples of interactions/moderators, interactions for continuous and categorical variables, separate models, fitted values <p>Reading:</p> <ul style="list-style-type: none"> • Allison: 166-171 • Hoffman pp. 17-19 • Kohler & Kreuter 220-223 • McLeod & Kaiser. 2004. Childhood emotional and behavioral problems and educational attainment. <i>American Sociological Review</i>, 69: 636-658. • Ellison, Christopher G., Jason D. Boardman, David R. Williams, and James Jackson. 2001. "Religious Participation and the Life-Stress Paradigm: Findings from the 1995 Detroit Area Study." <i>Social Forces</i> 80:215-49. <p>Assignment (due 3/13 in class):</p> <ul style="list-style-type: none"> • see handout.
3/13	<p style="text-align: center;">Generalized Linear Models</p> <p>Topics:</p> <ul style="list-style-type: none"> • The logic of the model, estimating the model <p>Readings:</p> <ul style="list-style-type: none"> • Hoffman: Ch 2 • Allison Ch. 9 <p>Assignment (due 3/20 in class):</p> <ul style="list-style-type: none"> • Handout and <u>draft of research questions and hypotheses.</u>
3/20	<p style="text-align: center;">Linear Probability, Logistic, Probit Regression</p> <p>Topics:</p> <ul style="list-style-type: none"> • Model specification, interpreting the results <p>Reading:</p> <ul style="list-style-type: none"> • Kohler & Kreuter 246-249 • Hoffman Ch. 3 • Boardman, Jason D. 2004. "Health pessimism among black and white adults: The role of interpersonal and institutional maltreatment." <i>Social Science & Medicine</i> 59: 2523-2533. <p>Assignment (due in class 4/3):</p> <ul style="list-style-type: none"> • see handout.
3/27	NO CLASS-Spring Break

4/3	<p style="text-align: center;">Ordered Logistic and Ordered Probit Regression</p> <p>Topics:</p> <ul style="list-style-type: none"> • Model specification, interpreting the results <p>Reading:</p> <ul style="list-style-type: none"> • Hoffman Ch. 4 • Keene & Reynolds. 2005. The Job Costs of Family Demands: Gender Differences in Negative Family-to-Work Spillover. <i>Journal of Family Issues</i>, 26 (3):275-299 <p>Assignment (due in class 4/10):</p> <ul style="list-style-type: none"> • <u>Draft of the Methods section of your paper.</u> This should include: a) Data; b) Measures; and c) Statistical analyses/plan.
4/10	<p style="text-align: center;">Multinomial Logistic Regression Model</p> <p>Topics:</p> <ul style="list-style-type: none"> • Model specification, interpreting the results <p>Reading:</p> <ul style="list-style-type: none"> • Hoffman Ch. 5 • Juby et al. 2005. Sharing roles, sharing custody? Couples' characteristics and children's living arrangements at separation. <i>Journal of Marriage and Family</i> 2005 67:1 157 <p>Assignment (due in class 4/17):</p> <ul style="list-style-type: none"> • see handout.
4/17	<p style="text-align: center;">Poisson and Negative Binomial Regression Models</p> <p>Topics:</p> <ul style="list-style-type: none"> • Model specification • Interpreting the results <p>Reading:</p> <ul style="list-style-type: none"> • Hoffman Ch. 6 • Skarupski et al. 2005. Black-White Differences in Depressive Symptoms Among Older Adults Over Time. <i>The Journals of Gerontology Series B: Psychological Sciences and Social Sciences</i> 60:P136-P142 <p>Assignment (due in class 4/24):</p> <ul style="list-style-type: none"> • <u>Provide a draft of the findings section of your paper.</u> This should include at least two tables: a) Descriptive statistics: provide the mean/s.d. (or %/N) for all variables used in your analyses; and b) Regression model estimates.
4/24	<p style="text-align: center;">Additional considerations and subsequent models</p> <p>Topics:</p> <ul style="list-style-type: none"> • Spatial analysis, multilevel analysis, growth models, event history models <p>Reading:</p> <ul style="list-style-type: none"> • Hoffman Ch. 8 <p>Assignment:</p> <ul style="list-style-type: none"> • No assignment this week
5/1	Individual Meetings
5/10	Final papers are due by 12 pm (noon) on Wednesday, May 10 th .